TÓM LƯỢC BÀI GIẢNG

(Vũ Quốc Hoàng)

ÔN TẬP CÁC KIỂU DỮ LIỆU CƠ BẢN

Chủ đề

- Kiểu dữ liệu
- Các kiểu dữ liệu cơ bản (trong C)

Tài liệu

[1] Vũ Quốc Hoàng, *Bí kíp luyện Lập trình C (Quyển 1)*, hBook, 2017.

Đọc tài liệu

- Kiểu dữ liệu và các kiểu dữ liệu cơ bản: Bài 3.5 [1]

Kiến thức

- Mô hình cơ bản của chương trình: nhập dữ liệu (input), xử lý dữ liệu (process), xuất dữ liệu (output). Theo đó, *dữ liệu* (data) đóng vai trò <u>trung tâm</u>.
- *Tháp dữ liệu* (DIKW pyramid) mô hình quan hệ <u>phân tầng</u> giữa dữ liệu với các đối tượng liên quan khác: *dữ liệu* (data), *thông tin* (information), *tri thức* (knowledge), *học thức* (wisdom). Thậm chí, *tín hiệu* (signal) có thể xem là tầng thấp nhất. Tuy nhiên, trong ngữ cảnh của môn học này, tất cả được gọi chung là dữ liệu.
- *Kiểu dữ liệu* (data type) là nhóm các dữ liệu giống nhau về ý nghĩa, mục đích sử dụng, dạng thức. Các dữ liệu của cùng một kiểu được tổ chức và xử lý giống nhau trên máy, trong chương trình.
- Một kiểu dữ liệu được xác định bởi: tập các dữ liệu cụ thể có thể có của kiểu, tập các thao tác xử lý có thể dùng trên dữ liệu của kiểu, cách định dạng và lưu trữ/truyền nhận dữ liệu của kiểu trên máy.
- Ở mức thấp nhất, dữ liệu là dãy *bit* (binary digit), được gom thành *byte*, *word*. Các bit, byte, word được tổ chức thành các dạng dữ liệu khác nhau. Có thể nói rằng:

  dữ liệu = bits/bytes/words + cách tổ chức/diễn giải/ngữ cảnh
- Các *kiểu dữ liệu cơ bản* (basic data type) là các kiểu dữ liệu đơn giản, dùng để xây dựng các kiểu dữ liệu khác. Trong C đó là các kiểu gần gũi với máy: *số nguyên* (integer), *số thực* (real), *luận lý* (boolean) và *kí tự* (character). Các kiểu này thường có *kích thước* (size), tức là số byte dùng để biểu diễn, cố định và nhỏ. *Chuỗi* (string) và *con trỏ* (pointer) cũng có thể xem là các kiểu dữ liệu cơ bản. (Chuỗi đã học còn con trỏ sẽ học sau.)
- Kiểu số nguyên thường được biểu diễn bằng dạng *bù 2* (two's complement) và được xác định cụ thể bởi kích thước và việc *có dấu hay không dấu* (signed/unsigned). Từ đó, *phạm vi* (range) của kiểu cũng được xác định.

- Kiểu số thực thường được biểu diễn bằng dạng *dấu chấm động* (floating point), cụ thể là *IEEE 754*, và được qui định bởi kích thước. Từ đó, phạm vi và *độ chính xác* (precision) của kiểu cũng được xác định.
- Giá trị luận lý thường được biểu diễn bằng kiểu số với qui ước 0 là *sai* (false) và khác 0 là *đúng* (true). Các thao tác cơ bản là: *và* (and), *hoặc* (or), *phủ định* (not). Giá trị luận lý thường được dùng trong các biểu thức điều kiện, kiểm tra. C++ có kiểu luận lý `bool` nhưng thật ra được biểu diễn bằng kiểu số nguyên.
- Kí tự thường được biểu diễn bằng số nguyên qua *mã ASCII* hay *mã Unicode*. Kí tự được dùng để xây dựng kiểu chuỗi.
- C/C++ là ngôn ngữ *định kiểu mạnh* và *tĩnh* (static and strongly typed): tất cả dữ liệu đều phải xác định kiểu lúc biên dịch. Kiểu dữ liệu của biểu thức được suy từ kiểu của các toán hạng và ngữ nghĩa của phép toán tương ứng. Các *hằng* (literal) cũng có kiểu tương ứng.
- *Chuyển kiểu* (type conversion) là việc chuyển một giá trị (dữ liệu) từ kiểu này (*kiểu nguồn*) sang kiểu khác (*kiểu đích*). Chuyển kiểu có thể là *ngầm định* (implicit) do C/C++ tự động thực hiện hay *tường minh* (explicit) do lập trình viên yêu cầu bằng *phép ép kiểu* (casting). Trường hợp chuyển kiểu ngầm định hay gặp trong các kiểu số là *phép nâng kiểu* (coercion).
- Kiểu dữ liệu được định danh bởi *tên kiểu*. C/C++ cho phép đặt nhiều tên khác nhau cho cùng một kiểu bằng `typedef`.

Kĩ năng

- Biết cách khai báo, thao tác với dữ liệu của các kiểu cơ bản
- Biết cách tra cứu và vận dụng các hàm trong thư viện chuẩn của C cho các thao tác liên quan đến các kiểu cơ bản (như `math.h` cho các thao tác trên số, `ctype.h` cho các thao tác trên kí tự, …)
- Viết được các hàm xử lý cho các thao trên các dữ liệu cơ bản
- Hiểu, nhận diện và vận dụng được các phép chuyển kiểu
- Vận dụng được các kiểu dữ liệu cơ bản để giải các bài toán thông dụng

Lưu ý

- Vì các kiểu dữ liệu cơ bản là nền tảng để xây dựng các kiểu dữ liệu khác nên sinh viên <u>cần phải</u> nắm vững và thành thạo
- Trong đa số trường hợp, sinh viên không cần nắm chi tiết biểu diễn/định dạng mức máy của các kiểu dữ liệu cơ bản

Bài tập

1. Tra cứu thêm về biểu diễn số nguyên dạng bù 2, số thực chấm động IEEE 754, các bảng mã ASCII, Unicode
2. Các bài tập 3.5 [1]